Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)

Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

A revised version of this paper is forthcoming in the American Journal of Agricultural Economics (AJAE)

Nr. 214

Revisiting the Role of Education for Agricultural Productivity

Malte Reimers, Stephan Klasen

November 2011

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173

e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

Revisiting the Role of Education for Agricultural Productivity

Malte Reimers[a*] and Stephan Klasen[a]

November 2011 (revised version of the Discussion Paper No. 90 of the Courant Research Centre "Poverty, Equity and Growth")

Abstract

While the majority of micro studies finds that rural education increases agricultural productivity, various recent cross-country regressions analyzing the determinants of agricultural productivity were only able to detect an insignificant or even surprisingly negative effect of schooling. In this paper, we show that this failure to find a positive impact of education in the international context appears to be a data problem related to the inappropriate use of enrolment and literacy indicators. Using a panel of 95 developing and middle-income countries from 1961 to 2002 that includes data on educational attainment, we show that education indeed has a highly significant, positive effect on agricultural productivity which is robust to the use of different control variables, databases and econometric methods. Distinguishing between different levels of education further reveals that only primary and secondary schooling attainment has a significant positive impact while the effect of tertiary education is insignificant. When distinguishing between income groups, our results indicate that even though the coefficient of the education variable is highly significant and positive for all quintiles, the returns to education are higher for the countries belonging to the richest three quintiles. This finding can be interpreted as support for the claim that education will have larger impacts on agricultural productivity in the presence of rapid technical change since it helps farmers to adjust more readily to the new opportunities provided by technological innovations.

Keywords: Agricultural productivity, agricultural production function, cross-country regression**,** education, human capital

JEL classification: I20, I25, O13, O15, O47, Q10

[a] University of Goettingen, Department of Economics, Goettingen, Germany. We would like to thank Summer Allen, Ousmane Badiane, Jacqueline Du Bois, Niels Kemper, Stefan Klonner, Sunday Odjo, Matin Qaim, Kimseyinga Savadogo, Matthias Schündeln, Susan Steiner, Michèle Tertilt, Maximo Torero, John Ulimwengu, and Fleur Wouterse for very helpful discussions and comments on earlier versions of this paper. Furthermore, we are grateful to all participants of the 2011 Convergence Project Workshop in Dakar, the Fourth Inter-University Doctoral Workshop in Development Economics in Heidelberg, and the 13th Annual Conference of the International Network for Economic Research (INFER) in London. This research is supported by the German Federal Ministry for Economic Cooperation and Development.
*Corresponding author: Malte Reimers; email: mreimer@uni-goettingen.de

1. Introduction

Early studies on the determinants of agricultural productivity across countries typically found significant positive coefficients for the education variables implying that higher levels of schooling lead to higher productivity (e.g. Hayami and Ruttan 1970; Nguyen 1979; Kawagoe et al. 1985; Lau and Yotopoulos 1989; Fulginiti and Perrin 1993). However, these findings are contrary to the ones of some newer studies applying more sophisticated econometric methods which either did not include education variables at all in the model (e.g. Frisvold and Ingram 1995), or found insignificant (Vollrath 2007) or even puzzling negative coefficients for the education variables used (Craig et al. 1997). Hence, the literature so far can be judged as rather inconclusive about the role of education for agricultural productivity in the international context.[1]

This is surprising given that the majority of micro studies finds a significant positive effect for education (e.g. Ali and Flinn 1989; Young and Deng 1999; Alene and Manyong 2006). For example, Philips (1994, p. 149) even states that "there is a general consensus that education has a positive effect on agricultural productivity". Regarding the mechanism leading education to affect agricultural productivity, various arguments have been proposed and empirically tested in the literature. First, education is supposed to let farmers become better "managers" by enhancing their decision-making skills (Asadullah and Rahman 2009). Second, education improves the peasant's access to information and therefore allows him/her to potentially pay and receive better prices for the inputs used and the outputs sold (Jamison and Lau 1982). Third, various empirical studies have shown that better educated farmers are adopting promising new technologies on average faster and therefore have a first mover advantage (Feder et al. 1985; Hossain et al. 1990; Lin 1991; Asfaw and Admassie 2004; Weir and Knight 2004). Lastly, it is regularly argued that as a consequence of the improved decision-making skills, better educated peasants are generally preferring riskier production technologies (typically promising higher returns) since they are able to evaluate adequately the implied opportunities and risks (Asadullah and Rahman 2009).

Given the preceding list of arguments supporting the view that rural education should enhance agricultural productivity, it remains an open question why cross-country studies using advanced econometric techniques were not able to find such an effect. In this paper,

[1] There is a related but somewhat distinct debate on the role of education on income growth in developing countries. See, for example, Pritchett (2001).

we show that these studies fail to detect the expected impact because they are using inadequate variables (enrolment and literacy rates) to approximate the stock of education. Using a panel of 95 developing and middle-income countries from 1961 to 2002 together with the newest version of the Barro-Lee educational attainment dataset (Barro and Lee 2010), we show that there is indeed a positive impact of educational attainment on agricultural productivity worldwide which is robust to changes in the control variables and in the econometric methods applied. Furthermore, distinguishing between different levels of education reveals that only primary and secondary schooling attainments have significant positive impacts on agricultural productivity. In addition, the prominent argument claiming that education leads to higher agricultural productivity particularly in the presence of rapid technical progress (Nelson and Phelps 1966; Schultz 1975; Rosenzweig 1995; Foster and Rosenzweig 1996) is tested empirically in the cross-country framework. Support is found indicating that the returns to education (in terms of augmentations of the agricultural productivity) are higher for countries with higher levels of income. We not only show these effects using our data set, but are also able to show that the inclusion of our measure of educational attainment is robust to the use of data sets from other studies that previously had found no impact of education. Furthermore, we submit our results extensive robustness checks and find very robust results in terms of magnitude and significance of effects.

The paper proceeds as follows. Section 2 gives an overview of the literature on education and its effects on agricultural productivity. Section 3 describes the indicators previous cross-country studies have included in their regression to control for education and argues why the average years of schooling as obtained from Barro and Lee (2010) are a conceptually superior proxy. Section 4 provides a description of the methodological approach and the data used. In section 5, the results of the fixed and random effects as well as feasible generalized least squares (FGLS) models are discussed and some further extensions of the models are introduced. Section 6 shows the results of diverse robustness checks. Finally, in section 7, the main results are summarized and conclusions are drawn.

2. Schooling and agricultural productivity: mechanisms and micro evidence

Before starting to systematically review the existing literature dealing with the question why increased schooling could have a positive impact on agricultural productivity, it is necessary to discuss what is considered to be the effect of education in general. Nelson and Phelps

(1966, p. 69) provide a widely-cited, relatively simple answer to this question by stating that education "enhances one's ability to receive, decode, and understand information". In addition, Schultz (1975, p. 835) argues that "education – even primary schooling – enhances the ability of students to perceive new classes of problems, to clarify such problems, and to learn ways of solving them". Welch (1970, p. 42) related the effect of education to agricultural production and identified two distinct phenomena through which schooling can have a productive value, namely the "worker effect"[2] and the "allocative effect". According to him, the former describes the phenomenon that well educated workers are simply able to use a given amount of resources more efficiently. In contrast, the latter is characterized by the ability of an educated worker to sufficiently "acquire and decode information about costs and productive characteristics of other inputs" (Welch 1970, p. 42). As a result, the highly educated farmer will regularly use a different mix of inputs compared to a relatively low-skilled peasant, i.e. he allocates his resources more efficiently. This phenomenon can hence be called the "allocative effect". With regard to the relevance of these two effects for agriculture, there is nowadays a consensus among researchers that farmer's schooling obtains its productive value mainly as a consequence of the allocative effect and only to a relatively limited extent from the worker effect (Huffman 1999).

Given that the concepts of the two above-described effects are still relatively vague, more recent literature has tried to further clarify the (often interrelated) transmission channels through which rural education may enhance agricultural productivity. However, it is important to emphasize that the basic notion of education as provided by Nelson and Phelps (1966) and Schultz (1975) is still relevant for these studies. The first argument commonly used to justify a potentially positive impact of education on agricultural productivity is a direct consequence of the above definitions. If one accepts that education allows farmers to make better use of the information available, to perceive new classes of problems and to find autonomous solutions to them, it directly follows that those peasants will possess superior decision-making skills and will hence be better "managers" allocating their resources more efficiently (Asadullah and Rahman 2009).

As a second argument, it is often claimed that well-educated farmers are not only capable of using available information more competently but also that they have better access to

[2] The phenomenon that Welch (1970) labeled as the "worker effect" is conceptually almost equivalent to what more recent studies typically describe with the term "technical efficiency" (Azhar 1991).

required information. Against the background that in many developing countries the majority of farmers have not received any schooling and are hence illiterate, it is easy to imagine that this lack of education is a severe obstacle for those peasants when seeking information. Thus, education provides peasants with the ability to disengage themselves from the "tight grip of […] inefficient 'word-of-mouth' communication patterns" (Welch 1970; Asfaw and Admassie 2004, p. 216). Taking this argument a step further, Jamison and Lau (1982) even argue that well educated farmers potentially pay and receive better prices for their inputs and outputs indicating that education can be a remedy to prevailing information asymmetries in the market.

A third argument that has received considerable attention in the literature suggests that well educated peasants are more likely to adopt new technologies or products early since they have superior access to related information and are better capable to distinguish between promising and unpromising innovations. In contrast, farmers with little schooling will often prefer not to introduce a new technology until its profitability has been proven e.g. until other farmers have successfully adopted the innovation (Nelson and Phelps 1966). This provides the educated farmer with a first-mover advantage making the new technology even more profitable and thus attractive for adoption. This argument is in line with the seminal contribution of Schultz (1975) who postulated that traditional agricultural societies are – in the absence of modernization – in an economic equilibrium since they have continuously optimized the use of the available resources over the generations. The occurrence of (exogenous) technological progress then pushes those societies away from the stationary state and allows them to achieve a superior equilibrium. However, the adjustment process takes time and its duration depends crucially on the population's ability to respond efficiently to the prevailing disequilibrium which can – according to Schultz (1975) – be enhanced through education. Consequently, the returns to education are expected to be higher in societies experiencing greater technical progress due to the henceforth increased level of complexity involved in the production process. On the contrary, in very traditional agricultural settings where tasks are rather simple, one would expect schooling to have only minor impacts on productivity (Schultz 1975; Schultz 1981; Rosenzweig 1995; Yang 1997). To the extent that income levels are correlated with levels and change of technology, this argument would suggest education to have a larger impact on agricultural productivity in richer countries.

In recent years, various authors empirically tested the above-described argument that educated farmers are more likely to adopt new technologies early and found overwhelming support for its validity (e.g. Feder et al. 1985; Hossain et al. 1990; Lin 1991; Asfaw and Admassie 2004; Weir and Knight 2004). Most prominently, Foster and Rosenzweig (1996) analyzed data from the green-revolution period in India and found increasing positive returns to schooling during this time of rapid technological progress. As an extension, they were even able to estimate the returns to schooling separately for areas faced with different levels of technological change.[3] The obtained results were in line with the above-described theoretical considerations from Schulz (1975) and a theoretical model provided by Rosenzweig (1995) indicating that the returns to schooling increased significantly more in those areas where a high degree of technological progress took place.

In recent years, a fourth argument has emerged from the literature that is strongly linked to the previous one. It is claimed that educated farmers adopt new technologies earlier because they tend to prefer riskier production technologies in general if these technologies provide higher expected returns (Knight et al. 2003; Asadullah and Rahman 2009). Hence, education is supposed to decrease the level of uncertainty and therefore to reduce the farmer's aversion towards endogenous risk[4], i.e. risks arising from the peasant's choice of production technology. Knight et al. (2003) tested this hypothesis empirically using household data from rural Ethiopia and found a significant reduction of risk aversion if the household head had received at least some schooling. This result implies that providing education to farmers not only lets them adopt *new* technologies earlier, but it may also change their attitude towards relatively risky *traditional* production technologies (e.g. crops they did not dare to plant before). As a consequence, the farmer may – after having received some schooling – optimize the mix of crops that he is planting (in the sense that he is now preferring riskier crops) since he is now better able to evaluate the associated risks and opportunities.

Besides the above-described literature, various relatively recent studies have focused on particular aspects of the relationship between education and agricultural productivity. For

[3] This extension was possible since the contemplated Indian areas differed substantially with regard to (exogenous) weather and soil conditions and therefore did not have the same ability to exploit the new seeds profitably (Foster and Rosenzweig 1996, p. 932).
[4] The literature typically subdivides risks into endogenous and exogenous ones (Knight et al. 2003). While the former arises directly from the peasant's choice of technology (e.g. mix of crops), the farmer is not able to influence the latter risk which is thus exogenous to him (e.g. variability in rainfall).

instance, Asadullah and Rahman (2009) distinguished in their analysis between different levels of education obtained by rice producing households in Bangladesh. Their results show – not entirely surprisingly – that basic education (defined as primary and secondary schooling) is relatively more important for agricultural productivity than higher education. Other studies have focused on the question whether agricultural households are benefitting from externality effects. While the findings consistently affirm the existence of intra-household externalities (Yang 1997; Asfaw and Admassie 2004) implying that there is indeed a knowledge-spillover among different members of agricultural households (even when the educated person is not involved in farming activities), the evidence for extra-household externalities (meaning that better educated neighbors have a positive impact on a household's productivity) is rather ambiguous (Foster and Rosenzweig 1995; Knight et al. 2003; Asadullah and Rahman 2009).

Given the highly aggregated nature of our country-level data, it is not possible to test adequately for the existence of externalities. However, we will make efforts to distinguish between different levels of education in our analysis in section 5.

Clearly, the literature demonstrates a range of plausible mechanisms linking education to higher agricultural productivity and these mechanisms are often found to empirically play a role in micro studies.

3. Problems and issues with measuring education in macro studies

In the macro-level cross-country and panel literature on the determinants of agricultural productivity worldwide, there are in total four human capital measures which are regularly included in the production function to account for differences in the quality of labor. Out of these four, every author typically uses two indicators in each model: one is to allow for differences in farmers' health (either life expectancy at birth or the total fertility rate) and one is to control for differentials in education (either the adult literacy rate or the gross/net enrolment ratio for primary/secondary education). Given the focus of this paper, we will in the following only discuss the appropriateness of the two education measures.

Despite their wide-spread use in the literature (e.g. Hayami and Ruttan 1970; Nguyen 1979; Kawagoe et al. 1985; Lau and Yotopoulos 1989; Fulginiti and Perrin 1993; Vollrath 2007[5]) both, gross and net enrolment ratios (GER and NER, respectively)[6], are rather inappropriate indicators for the current level of schooling in a country.[7] First, the data quality is often relatively poor since the enrolment rates are typically obtained from administrative records from schools which have a strong incentive to overstate the number of students in order to receive more resources for their institution. Second, the enrolment rates usually reflect the number of registered students at the beginning of the school year and thus do not take into account how many pupils drop out of class in the course of the year, i.e. they fail to adequately capture actual school attendance. Third and most importantly, enrolment ratios by definition just measure the *flow* of schooling and provide therefore information about the *future* and not the *current stock* of education in a country (Barro and Lee 1993). Only in the very particular case of constant enrolment rates for all countries, they would be able to mirror the steady-state stock of education correctly. However, this assumption is rather implausible given the substantial but heterogeneous increases in developing countries' enrolment ratios in recent years (Schultz 1988; Pritchett 2001). Hence, enrolment ratios – being gross or net – do not adequately reflect what the productivity literature wants them to reflect: the current stock of education available in a country.

As an alternative, the adult literacy rate, typically defined as the share of the population aged 15 and above having "the ability to read and write with understanding a simple statement related to one's daily life" (UNESCO 2011), has been used by various authors to approximate the population's level of schooling (e.g. Hayami and Ruttan 1970; Kawagoe et al. 1985; Lau and Yotopoulos 1989; Craig et al. 1997). This is not surprising given that the adult literacy rate possesses several features one would expect from a perfect measure for the level of education. First, the concept is relatively simple and the data are available for a wide array of countries. Second, the adult literacy rate indeed gives an idea of the current stock of education available in a country and is thus preferable to the enrolment ratios. However, there are also several drawbacks. Most importantly, the adult literacy rate must

[5] Vollrath (2007) states in footnote 5 of his paper that he tried to include primary enrolment ratios, but found them to be insignificant.
[6] According to the UNESCO (2011), the NER is defined as the "enrolment of the official age group for a given level of education expressed as a percentage of the corresponding population" as opposed to the GER being defined as the "total enrolment in a specific level of education, regardless of age, expressed as a percentage of the eligible official school-age population corresponding to the same level of education in a given school year".
[7] See also Barro and Lee (1993) for an extensive discussion of these issues.

be judged as a relatively "crude measure of schooling" (Huffman 1999, p. 31) since it just refers to the "first stage in the path of human capital formation" (Barro and Lee 1993, p. 367) and does hence not sufficiently allow to assess the full depth of education. As a consequence, the adult literacy rate – if used as an indicator for the quality of labor in a productivity analysis – has the inherent problem that it implicitly assumes that any education higher than the most elementary level will not have any productive value (Barro and Lee 1993). Furthermore, it is by definition bounded above implying that it is not possible to achieve literacy rates higher than 100 per cent. Because of this feature, the variation between countries is artificially reduced, in particular when contemplating middle- or high-income countries. These obvious drawbacks can, for instance, be exemplified with data for the Maldives and Israel. While the two countries are almost equal in terms of the adult literacy rate (97.0 and 97.1 per cent, respectively (UNDP 2009)), the average years of schooling differ substantially with a Maldivian adult having received on average 6.14 years of schooling compared to an average of 11.33 years for an Israeli (both numbers from Barro and Lee (2010)). This obvious discrepancy in the educational attainment in the two countries is not reflected sufficiently in the data and is thus ignored when taking the adult literacy rate to approximate the current stock of schooling.

Considering the problems of the two schooling indicators from an econometric point of view, one can consider enrolment and literacy rates as variables that measure the true stock of education with error. As is well-known, measurement error leads to a downward bias in estimated coefficients which might explain the failure to find effects using these proxies.

In short, the two above-described measures both suffer from severe methodological weaknesses and do not adequately reflect the stock of education currently available in a country. Against this background, Barro and Lee (1993) introduced already in the early nineties their educational attainment dataset which has since then been methodologically improved and regularly updated (Barro and Lee 1996, 2001, 2010). It is mostly based on reported school attainment data in census and household surveys (mainly compiled by UNESCO and Eurostat) which are then projected using robust simulation methods to generate the achievement data for the benchmark years. In particular, Barro and Lee (2010) calculate – as a first step – the educational attainment of the population by 5-year age groups and split then the distribution up into four rather broad attainment categories[8]. Forward and backward extrapolation is as a next step used to fill in missing observations with each group being assigned an age- and education-specific mortality rate (hence not assuming a uniform mortality). Nowadays, the variables from the dataset are widely accepted, not only in the economics literature, to presumably provide the most reliable proxies for the stock of education for a large number of countries.

For the analysis conducted in this paper, we therefore decided to use the newest version of their dataset offering 5-year-averages of the educational attainment for 146 countries reaching from the year 1950 to 2010. In particular, we will use data for the average number of years of schooling (s_t) for the population aged 15 and above which the two authors defined as

$$(1) \qquad s_t = \sum_{a=1}^{A} l_t^a \, s_t^a \qquad \text{with} \qquad s_t^a = \sum_j h_{j,t}^a \; Dur_{j,t}^{\,a}$$

where l_t^a denotes the share of age group *a* in the population aged 15 and above, s_t^a is the average number of years of schooling of age group *a*, h_j^a corresponds to the share of the age group *a* having attained the schooling level j = *primary, secondary, tertiary*, and *Dur* represents the duration in years corresponding to the respective level of education (Barro and Lee 2010, p. 7).

We argue that this indicator is methodologically superior to the measures previously used in the literature on the determinants of agricultural productivity worldwide because it shares the desirable characteristics of the adult literacy rate (relatively simple concept, good availability of data, actually measuring the *current* stock of education) and additionally has the advantage of not being restricted to the most basic level of education. Therefore, the variable accounts more adequately for the full depth of education.

However, it is important to emphasize that the Barro-Lee measures still do not meet two requirements that one would expect of the "perfect" measure of education in our particular context. First, the indicators are exclusively focused on the quantity of schooling received by the population and only partly account for quality differences. In particular, only to the extent that a student's achievements were insufficient to pass a grade will this be reflected in the educational attainment indicator which measures years passed (rather than years

[8] Namely: No formal education, primary education, secondary education, and tertiary education.

attended); quality differences beyond passing or failing a grade are not considered. Second, it would be highly desirable that the measure can be disaggregated into rural and urban areas since the vast majority of agricultural labor lives in rural areas. When testing the robustness of our results in section 6, we will make efforts to overcome this shortcoming. Despite these caveats, it is clearly the case that taking the average years of schooling as of Barro and Lee (2010) presents a crucial improvement to indicators previously used in the literature to approximate a country's stock of education.

From here onwards, the argument of this paper is as follows. Based on the extensive theoretical considerations provided in section 2 as well as the empirical evidence found in two early meta-studies (Lockheed et al. 1980; Phillips 1994[9]) and numerous micro studies (e.g. Ali and Flinn 1989; Young and Deng 1999; Alene and Manyong 2006) the hypothesis is that rural education indeed increases on average agricultural productivity. However, this stands in sharp contrast to recent cross-country studies applying modern econometric methods (particularly panel estimation techniques including time and country dummies) which either did not include any education variables in the model (e.g. Frisvold and Ingram 1995) or found insignificant (Vollrath 2007) or even puzzling negative coefficients for the education variables used (Craig et al. 1997). Against the background, of the above-described inadequateness of the education indicators used in those papers (adult literacy rate or gross/net enrolment ratio for primary/secondary education), we argue – in line with Huffman (1999, p. 31) – that the inability to detect the expected robust positive impact for education in the cross-country framework is rather due to data problems than due to the absence of real effects. This hypothesis will be tested in the following empirical part of the paper using the newest version of the Barro-Lee educational attainment dataset, but using the same advanced econometric framework and covariates of the recent studies that had failed to find an effect.

4. Methodology and data

The methodological approach applied in this paper is generally in line with the recent cross-country and panel literature on the determinants of agricultural productivity (e.g. Craig et al 1997, Vollrath 2007). We are assuming that the production process for the *i*th country at

[9] However, the results of such meta-studies should always be regarded with the necessary caution, since they implicitly assume that the methods and models of all contemplated studies were appropriate.

time t follows a common Cobb-Douglas production function. In particular, we estimate the following specification

$$(2) \qquad lny_{it} = \alpha + \beta_x lnX_{it} + \beta_E E_{it} + \beta_V V_{it} + \beta_C C_i + \gamma_i + \delta_t + \varepsilon_{it}$$

where the dependent variable is the natural logarithm of the output per ha and X_{it} is a vector of conventional agricultural inputs taken in per hectare terms. E_{it} is the above-described indicator for the average years of schooling as obtained from Barro and Lee (2010)[10]. Thus, β_E is the coefficient of main interest reflecting the partial productivity effect of education and it is expected to be positive. In addition, we also include V_{it} representing a vector of time-varying and C_i being a vector of time-invariant controls in the model. Lastly, γ_i and δ_t, respectively, are country- and time-specific constants typically included in panel models and ε_{it} is the potentially heteroskedastic error term.

Following standard practice in this literature,, we take the total value of all agricultural production after deductions for feed and seed (all expressed in 1999-2001 international $) divided by the total agricultural area in ha (both obtained from the FAOSTAT database) as the dependent variable (see Appendix A for the exact specifications and sources of all variables). The X_{it} vector contains four conventional inputs typically included in production functions, namely labor, fertilizer, tractors, and livestock (all in per hectare terms). Land is not included as a separate input in the equation since constant returns to scale are assumed and the variable thus cancels out. Data for the conventional inputs are all obtained from FAOSTAT whereat the livestock data is converted into cattle equivalents using weights from Hayami and Ruttan (1985)[11].

The vector of time-varying controls V_{it} can be subdivided into up to four categories. The first group of variables intends to account for differences in the quality of land. Therefore, we included the share of agricultural land equipped for irrigation and the percentage of agricultural land that is used as permanent meadows and pastures (both obtained from

[10] Against the background that the average years of schooling are only available in five year intervals while all other variables in our model are disposable on an annual basis, we decided to linearly interpolate the schooling data.

[11] Their widely-used weighting scheme allows to transform the headcounts of different animals into comparable units by assuming that 1 horse = 1 mule = 1 buffalo = 1.25 cattle = 1.25 asses = 0.9 camels = 5 pigs = 10 sheep = 10 goats = 100 chickens = 100 ducks = 100 geese = 100 turkeys.

FAOSTAT) into our regression.[12] To further allow for differentials in climate, we additionally used satellite data reflecting the average precipitation on agricultural land in year t for country i (data from Williams and Breneman (2009)). As a second group of time-varying controls, two road traffic-related variables are used to make sure that the human capital variables not just capture the potentially positive effect of a well developed infrastructure. Typically, the literature uses for this purpose the road density defined as the total road length in km per square km² of land area (data taken from Canning (1998) and WDI online, respectively). However, we argue that this concept is too narrow since it is rather not the pure disposability of roads that generates a productive value, but the effective ability of the economy to regularly use these roads. Therefore, we additionally try the per capita road sector energy consumption[13] (data from WDI online) to account for differentials in the infrastructure. Thirdly, it is important to rule out the possibility that discrepancies in the quality of institutions are driving the results. To allow for this, we further include the political risk index taken from the International Country Risk Guide (ICRG) which is a commonly used indicator for a country's political stability (Political Risk Services 2005). The last category contains additional human capital variables not being included in E_{it}. In particular, two alternative measures are, in line with the literature, used to account for differences in the population's health level, namely the life expectancy at birth and the total fertility rate. In addition to the above-described time-varying controls, dummy variables for the legal origin as derived by La Porta et al. (1999) were included in the C_i vector of our model to allow for time-invariant differences in the legal system of the countries.[14] To give an overview on the data used for the analysis, Table 1 presents summary statistics for the above-described variables.

[12] We also tried the land quality index of Peterson (1987) which is time-invariant and would thus belong to the vector C_i . However, it turned out that this variable greatly reduced our sample without adding any meaningful information.

[13] The correlation ρ between the two measures is in our dataset approx. 0.45 implying that it is in fact possible to include both variables at the same time without introducing a multicollinearity problem.

[14] According to La Porta et al. (1999), it is possible to classify a country's legal origin in one of the five following groups: English common law, French commercial code, German commercial code, Scandinavian commercial code, and Socialist/Communist law.

Table 1: Summary Statistics

Variable	Observations	Mean	SD	Min	Max
Net production per ha (intl. $)	3,282	271.87	296.28	2.68	2,063.74
Workers per 1,000 ha (number)	3,282	424.94	605.59	1.99	4,498.84
Tractors per 1,000 ha (number)	3,282	3.50	10.25	0.00	117.36
Livestock per 1,000 ha (in cow equivalents)	3,282	477.91	428.12	4.40	2,801.47
Fertilizer per 1,000 ha (in tons)	3,282	20.26	36.86	0.00	337.84
Irrigated land (in % of total)	3,282	6.26	10.10	0.00	67.12
Land in pasture (in % of total)	3,282	58.00	28.80	0.71	99.50
Precipitation on agricultural land (in mm)	2,184	1,164.65	701.55	6.00	3,738.00
Road density (km of roads per 100 sq. km of land)	1,654	17.24	24.32	0.26	220.13
Road sector energy consumption p. c. (kt of oil equivalent)	1,970	0.16	0.26	0.00	2.53
Life Expectancy at birth (in years)	3,282	57.98	10.16	26.41	78.88
Total fertility rate (children per women)	3,273	5.16	1.80	1.09	8.73
Total years of schooling	3,282	4.40	2.52	0.04	10.80
Years of primary schooling	3,282	3.13	1.68	0.04	7.31
Years of secondary schooling	3,282	1.15	1.00	0.00	5.64
Years of tertiary schooling	3,282	0.12	0.15	0.00	1.07
Political Risk Index (ICRG)	1,288	56.34	12.30	14.08	81.67

A basic assumption that is standard in cross-country regressions trying to explain differences in agricultural productivity is that there exists a common production function which is applicable to all countries in the sample – a so-called "meta-production function" (Hayami and Ruttan 1970; Kawagoe et al. 1985). Without any doubt, this assumption is strong and it can plausibly be argued that the agricultural production process differs between industrialized and developing countries.[15] Taking such objections serious, the currently 34 OECD member countries were dropped from the dataset in order to reduce its heterogeneity, so that we are left with a sample of developing and middle-income countries. We further excluded countries/territories with very small agricultural areas or labor forces to minimize measurement error since the corresponding data from FAOSTAT are generally rounded to thousands leading to a severe bias for countries having only small values for these variables.[16]

[15] This general drawback of cross-country regressions is sometimes ignored by researchers, leading to samples where the assumption of a common meta-production function is hardly defendable.

[16] Namely, the countries/territories dropped due to these two exclusion criteria are: American Samoa, Andorra, Anguilla, Aruba, Bahamas, Bahrain, Barbados, Bermuda, British Virgin Islands, Brunei, Cook Islands, Falkland Islands, Faroe Islands, French Guiana, Gibraltar, Greenland, Guadeloupe, Holy See, Kuwait, Liechtenstein, Luxembourg, Malta, Martinique, Monaco, Montserrat, Nauru, Netherlands Antilles, Niue, Norfolk Island, Northern Mariana Islands, Qatar, Palau, Saint Helena, Saint Kitts and Nevis, San Marino, Seychelles, Singapore, Tokelau, Turks and Caicos Islands, Tuvalu, United States Virgin Islands, Wallis and Futuna Islands.

Finally, we also made efforts to clean our sample of all observations that are biased due to major natural disasters and/or armed conflicts. To account for the former, we divided for each year the number of inhabitants affected by earthquake, floods or droughts (obtained from EM-DAT (2011)) by the total population of the country. All observations where this ratio exceeded the threshold of one third were then excluded from the analysis since we consider an efficient agricultural production under these circumstances as practically impossible. In addition, we took battle deaths data from the Centre for the Study of Civil War (Lacina and Gleditsch 2005) and analogously calculated the share of the population that was killed in the specific year due to armed conflicts. We argue that a share of 0.1% (i.e. one in one thousand inhabitants) together with the associated flow of refugees is sufficient to impede efficient agricultural production. Hence, we dropped all corresponding observations from the dataset (see Appendix B for a list of observations that were dropped due to the two exclusion criteria). As an alternative to dropping these observations, one could also include a dummy variable to account for natural disasters or armed conflicts. Not entirely surprising, the results for this second alternative do not differ from the ones obtained when dropping the affected observations.[17]

The result of these modifications is an unbalanced panel covering 95 countries for the time period from 1961 to 2002 (a detailed list of the countries and the number of observations is provided in Appendix C).

5. Regression results

As a first step of the analysis, two random-effect models (RE) are applied to the data. In the most basic specification (column 1), we include, besides the four conventional agricultural inputs (X_{it}), also the share of the agricultural land that is equipped for irrigation and the percentage used as permanent meadows and pastures. The coefficients of all variables are statistically significant and show the expected positive signs. The statistically significant negative coefficient for permanent meadows and pastures is also not surprising given that a high value for this indicator is typically a sign for a relatively low quality of the agricultural land which is presumably the reason for extensive land use as meadows and pastures.[18] We

[17] The results for this second alternative are available on request.

[18] This statement is supported by the fact that the variable is highly negatively correlated ($\rho \approx -0.75$) with the land quality index from Peterson (1987).

further include the ICRG political risk indicator (taken as an average for each country and assumed to be stable over time[19]) and the La Porta et al. (1999) legal origin dummies as two time-invariant country controls (C_i) as well as year dummies (δ_t). Of course, the variables of main interest are those measuring human capital, namely life expectancy at birth and particularly the average years of schooling, which both show highly significant and positive coefficients. This implies that a higher level of education increases agricultural productivity which is in line with the theoretical considerations from section 2. The point estimate suggests that an additional year of schooling improves agricultural productivity by around 6 percent, a sizable effect that is not only statistically but also economically significant.

[19] This assumption is in line with the literature (e.g. Vollrath 2007) and is in our case necessary since the political risk index is only available for the years 1984 onwards and we would thus lose all observations before this year. However, relaxing the assumption does not materially change our results (see robustness checks in Table 6).

Table 2: Results of the Panel Regressions

	(1) RE	(2) RE	(3) FE	(4) FE	(5) FGLS	(6) FGLS
(log) Livestock per ha	0.294***	0.291***	0.291***	0.293***	0.269***	0.291***
	(7.035)	(6.294)	(6.650)	(5.255)	(16.722)	(13.577)
(log) Fertilizer per ha	0.076***	0.056***	0.066***	0.060***	0.014***	0.017***
	(5.055)	(3.012)	(5.170)	(3.186)	(5.360)	(4.130)
(log) Tractors per ha	0.082***	0.095***	0.061***	0.076***	0.060***	0.063***
	(3.553)	(3.407)	(2.919)	(2.814)	(10.372)	(8.285)
(log) Workers per ha	0.162***	0.264***	0.106*	0.209***	0.177***	0.236***
	(2.830)	(4.281)	(1.738)	(2.766)	(8.023)	(8.067)
Area equipped for irrigation (%)	0.006**	0.004*	0.004*	0.004	0.005***	0.005***
	(2.264)	(1.918)	(1.691)	(1.579)	(5.493)	(4.650)
Permanent meadows and pastures (%)	-0.007**	-0.012***	-0.006*	-0.010***	-0.006***	-0.008***
	(-2.208)	(-4.557)	(-1.785)	(-2.670)	(-6.399)	(-7.068)
Life Expectancy at birth	0.010**	0.013***	0.011***	0.012***	0.011***	0.011***
	(2.549)	(3.428)	(3.030)	(3.173)	(8.814)	(7.239)
Total years of schooling	0.060**	0.065***	0.053**	0.063**	0.033***	0.032***
	(2.442)	(2.991)	(2.020)	(2.512)	(4.325)	(3.542)
Road sector energy consumption		0.254**		0.226*		0.264***
		(2.302)		(1.838)		(3.615)
(log) Precipitation (mm)		0.035		0.027		0.021**
		(1.268)		(1.161)		(2.394)
Constant	6.789***	5.555***	5.843***	5.874***	5.405***	5.561***
	(12.044)	(12.369)	(20.720)	(14.498)	(44.843)	(37.108)
Observations	2,791	1,609	3,282	1,685	3,282	1,685
Number of countries	79	69	95	74	95	74
Country controls included[a]	yes	yes	no	no	no	no
Time fixed effects	yes	yes	yes	yes	yes	yes
Country fixed effects	no	no	yes	yes	yes	yes
ε_{it} autocorrelation	none	none	none	none	AR(1)	AR(1)
Hausman test statistic[b]			175.09	68.57		
Hausman test p-value			0.00	0.00		
Wooldridge test statistic[c]			27.28	18.12		
Wooldridge test p-value			0.00	0.00		
R^2	0.891	0.936	0.840	0.912	n.a.	n.a.

Notes: The dependent variable is the logarithm of the net agricultural production per ha (in intl. $). Robust z-statistics are given in parentheses. Single asterisk (*) denotes significance at 10%, double asterisk (**) denotes significance at 5%, and triple asterisk (***) denotes significance at the 1% level.

[a] All random-effects specifications include dummies for legal origin from La Porta et al. (1999) and the ICRG-institutions index.

[b] Hausman test statistic is distributed as X^2_{42} in column (3) and X^2_{30} in column (4).

[c] Wooldridge test statistic is distributed as F(1,94) in column (3) and F(1,73) in column (4).

Until now, it could still be argued that this finding is rather a spurious correlation since no controls for differences in climate or infrastructure across the countries were included in the model. However, as can be seen in column 2, this is apparently not the case since the inclusion of corresponding variables[20], namely the road sector energy consumption and the natural logarithm of average precipitation, does not materially change the results for the schooling variable in terms of size and significance. Instead, this even increases its statistical

[20] Due to the poor availability of data, the inclusion of these two variables reduces the sample to 1,685 observations reaching from year 1976 to 2002.

significance to the 1% level and conspicuously enhances the explanatory power of our model (raising the R^2 from 0.891 to 0.936).

As a second step, we re-estimate the model using a fixed-effects specification (FE) with time dummies. When doing so, it is no longer possible to include the time-invariant controls (C_i) as separate variables because these would automatically be intercepted by the country-specific constant (γ_i). As can be seen in columns 3 and 4, using the fixed-effect instead of the random-effect specification does not materially change the results of our analysis. In particular, the coefficient of the schooling variable remains consistently positive, of similar size (5-6% return for a year of schooling) and highly significant. The Hausman specification test is then used to compare the FE-model with the RE-model implying that only the fixed-effect estimator is consistent and is thus preferable.

The two estimation methods used so far do not control for serial correlation in the error terms. In line with Vollrath (2007), this assumption is questionable in the context of agricultural productivity analysis since various types of shocks are probably persistent over time (e.g. adverse weather conditions). To account for this possibility, the Wooldridge test for serial correlation (see Wooldridge 2002, p. 282) is applied and in both cases, the hypothesis of no first-order autocorrelation is strongly rejected. Hence, it is necessary to allow ε_{it} not only to be heteroskedastic (by calculating robust standard errors) but also to permit the error structure to follow an AR(1) process of the type $\varepsilon_{it} = c + \rho\varepsilon_{i,t-1} + \eta_{it}$ with ρ having a value between 0 and 1 and η_{it} being a white noise process with zero mean and variance σ_η^2. With regard to the parameter ρ there are generally two possibilities. It could on the one hand be presumed that the errors follow a *unit-specific* first-order autoregressive process (thus using ρ_i instead of ρ in the equation). On the other hand, it is also possible to assume the parameter to be *homogenous* across countries (consequently using ρ). Beck and Katz (1995) convincingly showed – using Monte Carlo simulations – that the use of feasible generalized least squares (FGLS) under the assumption of a unit-specific ρ_i leads to severely underestimated standard errors, implying an extreme overconfidence in the coefficients, when T is not at least as large as N. Given that T in our dataset is considerably smaller than N (27 years compared to 95 countries), we decided to assume ρ to be homogenous across countries.

Hence, the two variants of the model are – as a third step – re-estimated using feasible generalized least squares methods with time and country dummies included in all specifications (columns 5 and 6). As can be seen, the FGLS results are generally in line with the ones obtained using the RE- and FE-models. However, the coefficients of some of the variables changed in magnitude and/or in statistical significance. Most notably, the t-value of the total years of schooling variable increases substantially when allowing for first-order auto-correlation while the absolute magnitude of the coefficient almost halves to 0.032 (column 6). Nevertheless, the impact of education on agricultural productivity is still sizeable implying that if each member of the population obtained an additional year of schooling, the agricultural productivity of the country would ceteris paribus increase by approx. 3.2%. To illustrate the economic relevance of the estimated effect, we also calculate the total contribution of the actually observed changes in the level of education to the observed changes in agricultural productivity. This is done by multiplying the total increase in the years of education between 1976 and 2002 with the estimated coefficient and dividing this product by the change in the log of the agricultural productivity between 1976 and 2002: $\frac{(6.48-3.55)\times 0.032}{(5.27-4.82)} \approx 20.84\%$. Using this approach, the change in the years of education accounted for more than 20% of the increase in agricultural productivity in the time period under investigation which is indeed a sizeable contribution.

As a next step, we try various extensions of our model (Table 3). Given that the Hausman specification test clearly negates consistency for all random effects estimations, we only show the FE and FGLS results whereat the FGLS estimates are preferable for the reasons discussed above. First, we additionally include two variables regularly used in the literature, namely the total fertility rate and the road density (e.g. Craig et al. 1997; Vollrath 2007). Both remain insignificant at all conventional levels while the coefficient of the average years of schooling remains relatively unaffected. Furthermore, the inclusion of these two additional controls greatly reduces our sample from 1,685 to only 737 observations. Consequently, we do not consider this extension as an improvement and therefore do not continue to include these two variables in our model. Second, we substitute our standard schooling variable by more disaggregated data reflecting the average years of schooling separately for primary, secondary and tertiary education (also obtained from Barro and Lee (2010)). The results indicate that the effect of an additional year of schooling conspicuously differs by type of education. In particular, we find in our preferred model (FGLS) the returns to primary and secondary education to be positive and statistically significant at the five per cent level whereas the effect of tertiary schooling on agricultural productivity is not significantly different from zero.

Table 3: Extensions of the Panel Model

	(1) FE	(2) FGLS	(3) FE	(4) FGLS	(5) FE	(6) FGLS
(log) Livestock per ha	0.304***	0.220***	0.293***	0.290***	0.271***	0.266***
	(3.248)	(6.531)	(5.047)	(13.481)	(4.821)	(12.234)
(log) Fertilizer per ha	0.054**	0.015**	0.060***	0.016***	0.049**	0.012***
	(2.241)	(2.573)	(3.252)	(4.020)	(2.498)	(3.249)
(log) Tractors per ha	0.107**	0.088***	0.076***	0.064***	0.068***	0.061***
	(2.659)	(6.812)	(2.843)	(8.311)	(2.684)	(7.984)
(log) Workers per ha	0.074	0.220***	0.216**	0.236***	0.339***	0.281***
	(0.488)	(4.353)	(2.524)	(7.758)	(4.060)	(9.057)
Area equipped for irrigation (%)	0.002	0.006***	0.003	0.005***	0.004**	0.005***
	(0.621)	(3.296)	(1.537)	(4.548)	(2.058)	(4.907)
Permanent meadows and pastures (%)	-0.011	-0.013***	-0.010***	-0.008***	-0.009**	-0.006***
	(-1.570)	(-5.973)	(-2.698)	(-6.952)	(-2.368)	(-5.419)
Life Expectancy at birth	0.013*	0.013***	0.012***	0.011***	0.010**	0.009***
	(1.879)	(4.020)	(3.101)	(7.222)	(2.157)	(6.235)
Total years of schooling	0.056*	0.048***				
	(1.755)	(3.251)				
Road sector energy consumption	0.302	0.310**	0.219*	0.272***	0.073	0.212**
	(1.556)	(2.532)	(1.841)	(3.705)	(0.197)	(2.427)
(log) Precipitation (mm)	0.031	0.024*	0.027	0.021**	0.025	0.023**
	(1.031)	(1.883)	(1.159)	(2.375)	(1.025)	(2.559)
Total fertility rate	-0.055	-0.000				
	(-1.228)	(-0.025)				
Road density	-0.001	0.000				
	(-0.997)	(0.951)				
Years of primary education			0.052	0.029**		
			(1.414)	(2.268)		
Years of secondary education			0.073	0.040**		
			(1.464)	(2.303)		
Years of tertiary education			0.110	-0.046		
			(0.432)	(-0.603)		
Income quintile 1 (poorest) * Schooling					0.025	0.021**
					(1.062)	(2.269)
Income quintile 2 * Schooling					0.039	0.028***
					(1.653)	(3.099)
Income quintile 3 * Schooling					0.055**	0.035***
					(2.386)	(3.915)
Income quintile 4 * Schooling					0.053**	0.032***
					(2.078)	(3.576)
Income quintile 5 (richest) * Schooling					0.060**	0.032***
					(2.332)	(3.556)
Constant	6.191***	6.185***	5.908***	5.557***	6.165***	5.568***
	(10.021)	(15.759)	(13.702)	(36.370)	(14.351)	(37.583)
Observations	737	736	1,685	1,685	1,556	1,556
Number of countries	57	56	74	74	73	73
Time fixed effects	yes	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes	yes
ε_{it} autocorrelation	none	AR(1)	none	AR(1)	none	AR(1)
Wooldridge test statistic[a]	8.97		18.12		15.28	
Wooldridge test p-value	0.00		0.00		0.00	
R^2	0.883	n.a.	0.913	n.a.	0.906	n.a.

Notes: The dependent variable is the logarithm of the net agricultural production per ha (in intl. $). Robust t-statistics are given in parentheses. Single asterisk (*) denotes significance at 10%, double asterisk (**) denotes significance at 5%, and triple asterisk (***) denotes significance at the 1% level.
[a] Wooldridge test statistic is distributed as F(1,48) in column (1), as F(1,73) in column (3), and F(1,72) in column (5).

With regard to the magnitude of the coefficients, it is a bit surprising that the coefficient for secondary education exceeds the one of primary schooling. However, one explanation for this finding could be that it is not just the pure ability to read and write causing the greatest impact on agricultural productivity, but advanced analytical skills (not provided in primary schools) which become – as extensively discussed in section 2 – particularly important when adopting new technologies.

As a third extension, we use the GDP per capita (PPP) from the Penn World Tables 7.0 (Heston et al. 2011) to subdivide our sample into five income quintiles (with quintile 1 being the poorest and quintile 5 the richest).[21] We then generate dummy variables for each income quintile and multiply those with the average years of schooling indicator. This allows us to estimate the effect of an additional year of schooling separately for the five income groups while – at the same time – maintaining the assumption of a common meta-production function for all countries. The aim of this exercise is to empirically test the above-described hypothesis that the returns to education are generally higher in those societies that experience greater technical progress since the involved tasks in such settings become more complex and thus require a higher level of education (Schultz 1975; Rosenzweig 1995; Foster and Rosenzweig 1996). Without any doubt, the GDP per capita is not a perfect measure for technical progress in an economy. However, we think it can plausibly be argued that the agricultural production process in richer societies is usually more modern and that the farmers in those countries typically have better access to technological innovations.

Our results (columns 5 and 6) generally confirm the predictions of the above-described hypothesis. In the fixed effects specifications, the coefficient of education is statistically

[21] The easiest way to do such a classification would be to simply take either the initial or average GDP per capita for each country for the whole time period and rank the countries accordingly. However, this procedure has the drawback that all observations of a country are assigned to exactly the same group what makes the results of the analysis very sensitive to the allocation of countries to the income groups. Therefore, we decided to pursue a slightly different, but methodologically probably superior approach by first subdividing our sample into five-year intervals and then using for each of the intervals the GDP per capita in the first year to assign the observations belonging to that five-year interval to one of the income groups. This procedure is repeated for all intervals and has the great advantage of assigning the countries to income quintiles more flexibly, thus allowing the countries to switch the income quintile in the course of time. This is in our opinion the most appropriate way to account for the very differential growth performances worldwide (compare e.g. Southeast Asia with Sub-Saharan Africa) that have increased the farmer's access to technological innovations in some countries more rapidly than in others. One could suspect that we may create an endogeneity problem when ranking the countries according to their GDP while using the net agricultural productivity per ha as our dependent variable since these two measures may be highly correlated. However, the correlation between these indicators is in fact relatively low ($\rho \approx 0.22$). In addition, we argue that if there was a bias due to endogeneity, it would skew the estimates for the poorest quintiles downwards and not upwards as countries with rising agricultural productivity are more likely to leave this quintile.

significant and positive only for the richest three quintiles while remaining statistically insignificant for the poorest two quintiles. In our preferred model (FGLS) the results are slightly different, indicating that the effect of an additional year of schooling is in fact highly significant and positive for all income quintiles. However, with regard to the magnitude of the coefficients, both models reveal a general trend of smaller coefficients for the schooling variable for poorer income quintiles. We interpret these results as support for the claim, already discussed in section 2, that in very traditional agricultural settings where tasks are typically rather simple, one would expect the returns to education to be smaller (Schultz 1975; Schultz 1981; Rosenzweig 1995; Yang 1997).

6. Robustness checks

As a first robustness check to our findings, we test in Table 4 the question whether it would have been possible to find the above-described positive impact of education on agricultural productivity when using – instead of the average years of schooling from Barro and Lee (2010) – the measures typically used in the literature to approximate the current stock of education in a country (namely gross/net enrolment ratios and adult literacy rates).

Table 4: Robustness Checks 1

	(1) FE	(2) FGLS	(3) FE	(4) FGLS	(5) FE	(6) FGLS
(log) Livestock per ha	0.293***	0.291***	0.341***	0.291***	0.316***	0.338***
	(5.255)	(13.577)	(5.252)	(13.652)	(4.498)	(14.053)
(log) Fertilizer per ha	0.060***	0.017***	0.062***	0.020***	0.062**	0.022***
	(3.186)	(4.130)	(3.422)	(4.754)	(2.551)	(4.344)
(log) Tractors per ha	0.076***	0.063***	0.059**	0.047***	0.053*	0.051***
	(2.814)	(8.285)	(2.147)	(5.677)	(1.822)	(6.873)
(log) Workers per ha	0.209***	0.236***	0.224***	0.261***	0.264**	0.254***
	(2.766)	(8.067)	(2.891)	(8.560)	(2.183)	(6.027)
Area equipped for irrigation (%)	0.004	0.005***	0.003	0.005***	0.004	0.003***
	(1.579)	(4.650)	(1.160)	(3.523)	(1.542)	(3.136)
Permanent meadows and pastures (%)	-0.010***	-0.008***	-0.011***	-0.008***	-0.010**	-0.007***
	(-2.670)	(-7.068)	(-2.926)	(-6.972)	(-2.135)	(-5.205)
Life Expectancy at birth	0.012***	0.011***	0.013***	0.011***	0.010*	0.010***
	(3.173)	(7.239)	(2.973)	(6.829)	(1.925)	(6.197)
Road sector energy consumption	0.226*	0.264***	0.221	0.274***	0.181	0.180**
	(1.838)	(3.615)	(1.445)	(3.714)	(1.161)	(2.164)
(log) Precipitation (mm)	0.027	0.021**	0.012	0.022**	0.006	0.018**
	(1.161)	(2.394)	(0.585)	(2.552)	(0.255)	(1.963)
Total years of schooling	0.063**	0.032***				
	(2.512)	(3.542)				
Gross enrolment ratio			-0.002**	-0.000		
			(-2.495)	(-0.561)		
Adult literacy rate					0.003	0.001
					(0.735)	(1.261)
Constant	5.874***	5.561***	6.453***	4.661***	6.238***	5.742***
	(14.498)	(37.108)	(14.972)	(29.045)	(9.925)	(34.886)
Observations	1,685	1,685	1,575	1,574	1,321	1,313
Number of countries	74	74	85	84	84	76
Time fixed effects	yes	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes	yes
ε_{it} autocorrelation	none	AR(1)	none	AR(1)	none	AR(1)
Wooldridge test statistic[a]	18.12		18.12		14.70	
Wooldridge test p-value	0.00		0.00		0.00	
R^2	0.912	n.a.	0.872	n.a.	0.894	n.a.

Notes: The dependent variable is the logarithm of the net agricultural production per ha (in intl. $). Robust t-statistics are given in parentheses. Single asterisk (*) denotes significance at 10%, double asterisk (**) denotes significance at 5%, and triple asterisk (***) denotes significance at the 1% level.
[a] Wooldridge test statistic is distributed as $F(1,73)$ in column (1), as $F(1,82)$ in column (3), and $F(1,72)$ in column (5).

While we find a highly significant, positive impact of education when using the average years of schooling (columns 1 and 2 – identical to columns 4 and 6 in Table 2), the results change conspicuously when we simply substitute this variable with the gross enrolment ratio (GER) for primary education (columns 3 and 4) or the adult literacy rate[22] (columns 5 and 6). In particular, when using the GER as a proxy for the current stock of education, the FE-model indicates a highly significant, negative effect of education. In contrast, we find the coefficient

[22] Given that the available adult literacy data include many gaps, we decided to linearly interpolate the existing data.

not to be significantly different from zero in the FGLS model.[23] Similarly, when taking the adult literacy rate, the coefficient for education remains insignificant at all conventional levels in both models. We interpret these results as strong support for our argument from section 3 claiming that both indicators suffer from severe methodological weaknesses and are therefore inadequate proxies for the current stock of education in a country. Nevertheless, it is striking that the use of these indicators can actually impede the detection of the – at least according to our analysis – existing positive effect of education on agricultural productivity.

Secondly, we want to test whether it would have been possible to find the above-described significant positive impact of education with datasets that previous authors have used to explain differences in agricultural productivity worldwide. Therefore, we took the dataset of Vollrath (2007) and exactly replicated the panel results of his analysis (see Table 5 columns 1 to 3 – corresponding to columns 4, 2 and 6 in Table 6 in his paper). We then re-estimated the model including the interpolated average years of schooling variable (everything else unchanged) which minimally reduced our sample due to the unavailability of education data. As can be seen, schooling has indeed not only in our sample but also in the one Vollrath had used a highly significant positive effect on agricultural productivity regardless which estimation technique is used (RE, FE or FGLS) and whether one further controls for the agricultural R&D expenditures per ha (see Table 5 column 7 – corresponding to column 7 in Table 6 in Vollrath's paper). In fact, including such an education indicator raises the explanatory power of the model (see the increases in the R^2 between columns 1 and 4 and columns 2 and 5, respectively). These results are particularly interesting given that Vollrath states in footnote 5 of his paper that he tried to include primary school enrolment rates as a proxy for the level of education which did not add any meaningful information to the regressions and were thus left out. Against the background of the methodological superiority of the Barro-Lee measure over the enrolment rates (see section 3) and the results in Table 4, the contrasting results are not surprising. Instead, we interpret the findings of this second robustness check as support for the claim that the results are not dependent on our particular dataset or the empirical methodology applied.

[23] We also used the net enrolment ratio (NER) for primary schooling instead of the average years of schooling. The results are very similar to using GER and available on request.

Table 5: Robustness Checks 2 (Replication of Vollrath (2007)) [a]

	(1) RE	(2) FE	(3) FGLS	(4) RE	(5) FE	(6) FGLS	(7) FGLS
Gini coefficient	-0.49***	-0.50***	-0.48***	-0.51***	-0.49***	-0.07	-0.47***
	(-4.58)	(-4.22)	(-4.82)	(-4.72)	(-4.01)	(-0.80)	(-6.44)
Log avg. farm size	0.02	0.02	-0.06***	0.02	0.03	-0.03	-0.00
	(1.30)	(1.23)	(-2.88)	(1.28)	(1.50)	(-1.46)	(-0.27)
Inputs							
Log livestock per ha	0.41***	0.39***	0.32***	0.42***	0.39***	0.32***	0.42***
	(11.13)	(9.76)	(12.18)	(11.67)	(9.35)	(12.59)	(20.80)
Log fertilizer per ha	0.04***	0.04***	0.01**	0.04***	0.04***	0.02***	0.01
	(4.03)	(4.13)	(2.27)	(4.09)	(4.17)	(3.53)	(1.47)
Log tractor per ha	0.03***	0.03***	0.07***	0.03***	0.03***	0.08***	0.09***
	(3.15)	(2.85)	(6.93)	(3.53)	(2.67)	(7.86)	(8.85)
Log labor per ha	0.09***	0.07**	0.26***	0.11***	0.07**	0.39***	0.30***
	(3.43)	(2.53)	(7.67)	(4.23)	(2.50)	(12.26)	(9.51)
Land quality							
% Irrigated	1.27***	1.37***	0.34**	1.23***	1.49***	0.17	0.51***
	(8.59)	(8.84)	(2.19)	(7.68)	(9.15)	(1.38)	(5.07)
% Permanent pasture	-0.17	0.24	-0.56***	-0.35**	0.19	-0.71***	-0.49***
	(-0.98)	(1.07)	(-6.37)	(-2.24)	(0.82)	(-7.77)	(-7.39)
Total years of schooling				0.03***	0.04***	0.13***	0.08***
				(3.18)	(3.00)	(12.30)	(7.36)
Research effort							
Log. agric. R&D expend per ha							0.04***
							(4.77)
Constant	2.52***	3.29***	2.94***	3.27***	3.51***	4.66***	3.92***
	(5.86)	(7.48)	(9.51)	(7.85)	(7.84)	(15.30)	(15.04)
Observations	1,159	1,159	1,159	1,128	1,128	1,128	993
Number of countries	54	54	54	52	52	52	42
Country controls (Z)[b] included	Yes	No	Yes	Yes	No	Yes	Yes
ε_{it} autocorrelation	none	none	AR(1)	none	none	AR(1)	AR(1)
R^2	0.864	0.749	n.a.	0.886	0.788	n.a.	n.a.

Notes: The dependent variable is the logarithm of the net agricultural production per ha (in intl. $). Robust t-statistics are given in parentheses. Single asterisk (*) denotes significance at 10%, double asterisk (**) denotes significance at 5%, and triple asterisk (***) denotes significance at the 1% level. This table was created using STATA Version 9 to achieve an exact replication of the results of Vollrath (2007).

[a] All specifications further include the total fertility rate, life expectancy, and year dummies.

[b] Z includes the Kaufmann et al. (2002) index of institutions, dummies for legal origin from La Porta et al. (1999), and the land quality index from Peterson (1987).

As a third robustness check, we relax the assumption commonly made in the literature of a stable institutions index (Table 6 - columns 1 and 2). Given that the ICRG political risk index is only available for the years 1984 onwards, this modification greatly reduces our sample to only 1,120 observations. It turns out that the political risk variable itself does not have any significant impact on agricultural productivity. In addition, the effect of the average years of schooling remains positive and statistically significant at the 5 percent level in our preferred model (FGLS) and only slightly misses significance in the FE specification.

Fourth, the model is re-estimated using five-year averages instead of annual data in order to minimize the effects of perennial temporary shocks (see Table 6 - columns 3 and 4). Yet, this modification alters only slightly the magnitude of the coefficients for the education variable, but does not affect its statistical significance.

Table 6: Robustness Checks 3

	(1) FE	(2) FGLS	(3) FE	(4) FGLS
	annual data		5-year averages	
(log) Livestock per ha	0.232***	0.243***	0.303***	0.295***
	(4.539)	(9.889)	(4.715)	(13.418)
(log) Fertilizer per ha	0.046**	0.021***	0.096***	0.084***
	(2.098)	(3.903)	(3.930)	(11.571)
(log) Tractors per ha	0.090***	0.064***	0.054**	0.047***
	(2.940)	(6.535)	(2.236)	(5.514)
(log) Workers per ha	0.262***	0.284***	0.197***	0.207***
	(3.229)	(7.710)	(2.711)	(8.307)
Area equipped for irrigation (%)	0.004*	0.002*	0.002	0.001
	(1.794)	(1.888)	(0.960)	(0.924)
Permanent meadows and pastures (%)	-0.009**	-0.005***	-0.005	-0.006***
	(-2.043)	(-3.665)	(-1.342)	(-4.246)
Life Expectancy at birth	0.012***	0.011***	0.011**	0.011***
	(3.122)	(6.681)	(2.565)	(10.452)
Road sector energy consumption	-0.060	0.064	0.212*	0.196**
	(-0.203)	(0.657)	(1.987)	(2.472)
(log) Precipitation (mm)	0.040	0.021**	-0.026	0.008
	(1.361)	(2.010)	(-0.351)	(0.219)
Total years of schooling	0.046	0.027**	0.056**	0.044***
	(1.525)	(2.341)	(2.097)	(5.135)
Political Risk Index	-0.000	0.000		
	(-0.020)	(0.643)		
Constant	5.800***	5.532***	6.089***	5.910***
	(11.921)	(29.443)	(9.216)	(20.474)
Observations	1,120	1,120	396	395
Number of countries	69	69	74	73
Time fixed effects	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes
ε_{it} autocorrelation	none	AR(1)	none	AR(1)
Wooldridge test statistic[a]	14.79		21.94	
Wooldridge test p-value	0.00		0.00	
R^2	0.902	n.a.	0.913	n.a.

Notes: The dependent variable is the logarithm of the net agricultural production per ha (in intl. $). Robust t-statistics are given in parentheses. Single asterisk (*) denotes significance at 10%, double asterisk (**) denotes significance at 5%, and triple asterisk (***) denotes significance at the 1% level.
[a] Wooldridge test statistic is distributed as F(1, 68) in column (1), and F(1,70) in column (3).

Fifth, it was argued earlier in the paper that the average years of schooling as provided by Barro and Lee (2010) are methodologically superior to the measures previously used in the literature to approximate for the current stock of education in a country. However, the indicator still has the disadvantage of not solely measuring the education of the rural population which would be highly desirable given that the vast majority of agricultural labor comes from pastoral surroundings. Ulubaşoğlu and Cardak (2007) made an effort to address this issue and combined data from the UNESCO Educational Yearbooks and the World Bank

Education Statistics in order to calculate the average years of schooling separately for urban and rural areas.[24] As a robustness check to our analysis, we take these data and use them to predict the average years of schooling for the rural population by first regressing the rural on the national years of education and its square (both from Ulubaşoğlu and Cardak (2007) - see Appendix D). The resulting coefficients are then used to predict the average years of schooling for the rural population for all countries of our sample. As an alternative approach, we regress – as a first step – the ratio of rural to urban years of education on the nation's average years of education (see Appendix E) and make then – as a second step – use of the formula $s_{rural} = \frac{s_{national}}{\frac{\omega_{urban}}{r} + \omega_{rural}}$ to predict the avg. years of education for the rural population (the derivation of this formula is provided in Appendix F). The resulting predicted data are highly correlated with the Barro-Lee indicator for the average years of education used in the main part of our analysis (ρ≈0.99 and ρ≈0.92, respectively) and it is therefore not entirely surprising that replacing the Barro-Lee measure with the predicted values for the rural population does not significantly change our results (see Appendix G).

In addition, we would have liked to account for differences in the quality of schooling in our analysis instead of solely focusing on its quantity. Eric A. Hanushek and Ludger Wößmann have worked extensively on this topic and compiled a dataset of test scores for approx. 50 countries worldwide (Hanushek and Wößmann 2007). However, this dataset has – at least from our perspective – the drawback that a large part of these countries are industrialized nations which we intentionally excluded from our sample (see section 4). In addition, it is – according to Hanushek and Wößmann (2007) – necessary to take an average of the test scores over at least the last 40 years, in order to obtain a reliable proxy for the educational performance of the entire labor force and not just a measure of the quality of current students. When doing so, one ends up with just one observation per country, thus having a time-invariant quality of schooling indicator which would in our preferred models be simply intercepted by the country fixed effect. In short, the unavailability of a time series schooling quality indicators unfortunately prevented us from accounting for differences in the quality of education in our analysis.[25]

[24] However, due to data limitations this was only possible for a relatively small sample (in total 76 observations from 56 countries).

[25] As an additional test, some previous authors included the natural logarithm of the agricultural R&D expenditures per hectare in their models to account for the country's research effort. Despite a very poor

Sixth, Schultz (1999) and Wouterse (2011) claimed that increased farmer's human capital will not instantaneously translate into higher agricultural productivity and it is therefore necessary to consider lagged values of those variables. We took these objections serious and included two-year-lags for the life expectancy and the average years of schooling instead of current values in our model (results not shown). However, it turned out that this does not materially affect the results of our analysis.

Lastly, it could be argued that we may have a spurious correlation problem since our time dimension is relatively large (T=27). To meet these objections, we conduct unit root tests to check whether our dependent variable is a non-stationary series, i.e. is integrated of order one. The results of the Fisher tests (as proposed by Maddala and Wu (1999)) are unambiguous: both, the augmented Dickey-Fuller as well as the Phillips-Perron test, clearly reject the null hypothesis of an existing unit root. Hence, spurious correlation should not influence inference in our case.

7. Conclusion

In this article, we re-examine the role of education for agricultural productivity in a cross-country framework. It was claimed that recent cross-country studies using sophisticated econometric methods failed to detect a statistically significant, positive impact of schooling as a consequence of inadequate proxies used to measure a country's stock of education. Using a large panel of 95 developing and middle-income countries ranging from the years 1961 to 2002 together with the newest version of the educational attainment dataset of Barro and Lee (2010), we find that education in fact has a significant positive impact on agricultural productivity worldwide which is robust to a multiplicity of changes to the model and its estimation. The effect is sizeable, implying that an additional year of schooling for the

---

availability of data, we did the same as a robustness check (results not shown) using data from the ISNAR Agricultural Research Indicator Series (Pardey and Roseboom 1989). While the magnitude of the coefficient for the average years of education remained relatively unchanged, the variable now missed statistical significance. However, we argue that this is rather due to the dramatically reduced sample of only 272 observations from 49 countries (compared to 1,685 observations from 74 countries before) than a consequence of the inclusion of the R&D expenditures variable which always remained statistically insignificant with t-values below 0.30. To support our claims of sample size problems, we can show that the schooling variable in this particular reduced subsample was not statistically significant even when applying our most basic regressions (without agricultural R&D expenditures) and the inclusion of the additional control variable did not materially change any of the coefficients. Hence, it is rather the smaller and apparently biased subsample that caused the education variable to be insignificant and not the effect of the additional control variable for R&D expenditures.

whole population would raise agricultural productivity by approx. 3.2 % in the preferred FGLS model. Furthermore, we find that only primary and secondary formation has a statistically significant positive impact on agricultural productivity. Finally, the effect of schooling was estimated separately for countries of different income levels. Our results suggest that the effect of education is generally smaller for the poorest countries. These findings are in line with the arguments proposed by Schultz (1975), Rosenzweig (1995), and Foster and Rosenzweig (1996) claiming that in very traditional agricultural settings where tasks are typically rather simple, one would expect the returns to education to be smaller (compared to countries facing rapid technical change).

The policy implications of our paper are relatively straight forward. The positive impact of schooling on agricultural productivity found in our analysis supports the view that education is indeed one of the key ingredients to enhance productivity in developing and middle-income countries. Hence, even governments of nations relying to a great extent on the primary sector should maximize efforts to increase the population's level of education. However, in particular for the poorest countries, our findings underline the complementarity of capital investments in the agricultural and the education sector since technical progress is needed to fully exploit the productivity-enhancing potential of schooling. Or to say it with the words of Foster and Rosenzweig (1996, p. 951):

> *"…the returns to investment in technical change will in general be higher when primary schooling is accessible and the returns to investment in schooling will be higher when technical change is more rapid."*

We conclude with some caveats and further suggestions. First of all, our results are based on cross-country regressions which ought to be seen with the necessary caution since they all rely on relatively strong assumptions (e.g. the existence of a common meta-production function). However, we did our best to reduce the heterogeneity of the sample and are therefore relatively confident that this assumption is in our case justifiable. The fact that our macro findings are much more in line with the micro literature than previous macro findings further supports our contention. Secondly, our results may actually underestimate the total impact of education. In line with the arguments by e.g. Huffman (1999), Yang and An (2002) as well as Joliffe (2004), we believe that highly skilled individuals will typically seek work in the non-farming sector where the returns to their knowledge are usually higher. Hence, the results of our analysis do obviously not reflect the "full" effect of education for a country because we only consider the impact on agricultural productivity. Future research might want to consider this externality more directly.

8. References

Alene, A. D., and V. M. Manyong. 2006. "Endogenous Technology Adoption and Household Food Security: The Case of Improved Cowpea Varieties in Northern Nigeria." *Quarterly Journal of International Agriculture*, *45*(3): 211-30.

Ali, M., and J. C. Flinn. 1989. "Profit Efficiency among Basmati Rice Producers in Pakistan Punjab." *American Journal of Agricultural Economics*, *71*(2): 303-10.

Asadullah, M. N., and S. Rahman. 2009. "Farm Productivity and Efficiency in Rural Bangladesh: The Role of Education Revisited." *Applied Economics*, *41*(1): 17-33.

Asfaw, A., and A. Admassie. 2004. "The Role of Education on the Adoption of Chemical Fertiliser under Different Socioeconomic Environments in Ethiopia." *Agricultural Economics*, *30*(3): 215-28.

Azhar, R. A. 1991. "Education and Technical Efficiency During the Green Revolution in Pakistan." *Economic Development and Cultural Change*, *39*(3): 651-65.

Barro, R. J., and J. W. Lee. 1993. "International Comparisons of Educational Attainment." *Journal of Monetary Economics*, *32*(3): 363-94.

Barro, R. J., and J. W. Lee. 1996. "International Measures of Schooling Years and Schooling Quality." *The American Economic Review*, *86*(2): 218-23.

Barro, R. J., and J. W. Lee. 2001. "International Data on Educational Attainment: Updates and Implications." *Oxford Economic Papers*, *53*(3): 541-63.

Barro, R. J., and J. W. Lee. 2010. "A New Data Set of Educational Attainment in the World, 1950–2010." *NBER Working Paper*.

Beck, N., and J. N. Katz. 1995. "What to Do (and Not to Do) with Time-Series Cross-Section Data." *The American Political Science Review*, *89*(3): 634-47.

Canning, D. 1998. "A Database of World Stocks of Infrastructure, 1950–95." *The World Bank Economic Review*, *12*(3): 529-47.

Craig, B. J., P. G. Pardey, and J. Roseboom. 1997. "International Productivity Patterns: Accounting for Input Quality, Infrastructure, and Research." *American Journal of Agricultural Economics*, *79*(4): 1064-76.

EM-DAT. 2011. "The OFDA CRED International Disaster Database." (June 3, 2011), *available at* www.emdat.be

Feder, G., R. E. Just, and D. Zilberman. 1985. "Adoption of Agricultural Innovations in Developing Countries: A Survey." *Economic Development and Cultural Change*, *33*(2): 255-98.

Foster, A. D., and M. R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy*, *103*(6): 1176-1209.

Foster, A. D., and M. R. Rosenzweig. 1996. "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution." *The American Economic Review*, *86*(4): 931-53.

Frisvold, G., and K. Ingram. 1995. "Sources of Agricultural Productivity Growth and Stagnation in Sub-Saharan Africa." *Agricultural Economics*, *13*(1): 51-61.

Fulginiti, L. E., and R. K. Perrin. 1993. "Prices and Productivity in Agriculture." *The Review of Economics and Statistics*, 471-82.

Hanushek, E., and L. Wößmann. 2007. "The Role of Education Quality for Economic Growth." *World Bank Policy Research Working Paper No. 4122*.

Hayami, Y., and V. W. Ruttan. 1970. "Agricultural Productivity Differences among Countries." *The American Economic Review*, *60*(5): 895-911.

Hayami, Y., and V. W. Ruttan. 1985. *Agricultural Development: An International Perspective*. Baltimore: Johns Hopkins University Press.

Heston, A., R. Summers, and B. Aten. 2011. "Penn World Table Version 7.0" Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. May 2011.

Hossain, M., M. A. Quasem, M. M. Akash, and M. A. Jabber. 1990. *Differential Impact of Modern Rice Technology: The Bangladesh Case*. Dhaka: Bangladesh Institute of Development Studies

Huffman, W. E. 1999. "Human Capital: Education and Agriculture," B. L. Gardner , and G. C. Rausser, *Handbook of Agricultural Economics.* Amsterdam: Elsevier Science, 333-81.

Jamison, D. T., and L. J. Lau. 1982. *Farmer Education and Farm Efficiency*. Baltimore: The Johns Hopkins University Press.

Jolliffe, D. 2004. "The Impact of Education in Rural Ghana: Examining Household Labor Allocation and Returns On and Off the Farm." *Journal of Development Economics*, *73*(1): 287-314.

Kaufmann, D., A. Kraay, and P. Zoido. 2002. "Governance Matters II: Updated Indicators for 2000-01." *World Bank Policy Research Department Working Paper*.

Kawagoe, T., Y. Hayami, and V. W. Ruttan. 1985. "The Intercountry Agricultural Production Function and Productivity Differences among Countries." *Journal of Development Economics*, *19*(1): 113-32.

Knight, J., S. Weir, and T. Woldehanna. 2003. "The Role of Education in Facilitating Risk-Taking and Innovation in Agriculture." *Journal of Development Studies*, *39*(6): 1-22.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, and Organization*, *15*(1): 222-79.

Lacina, B., and N. P. Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population*, *21*(2-3): 145-66.

Lau, L. J., and P. A. Yotopoulos. 1989. "The Meta-Production Function Approach to Technological Change in World Agriculture." *Journal of Development Economics*, *31*(2): 241-69.

Lin, J. Y. 1991. "Education and Innovation Adoption in Agriculture: Evidence from Hybrid Rice in China." *American Journal of Agricultural Economics*, *73*(3): 713-23.

Lockheed, M. E., T. Jamison, and L. J. Lau. 1980. "Farmer Education and Farm Efficiency: A Survey." *Economic Development and Cultural Change*, *29*(1): 37-76.

Maddala, G. S., and S. Wu. 1999. "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test." *Oxford Bulletin of Economics and Statistics*, *61*(S1): 631-52.

Nelson, R. R., and E. S. Phelps. 1966. "Investment in Humans, Technological Diffusion, and Economic Growth." *The American Economic Review*, *56*(1/2): 69-75.

Nguyen, D. 1979. "On Agricultural Productivity Differences among Countries." *American Journal of Agricultural Economics*, *61*(3): 565-70.

Pardey, P. G., and J. Roseboom. 2004. *ISNAR Agricultural Research Indicator Series: A Global Data Base on National Agricultural Research Systems*. Cambridge: Cambridge University Press.

Peterson, W. 1987. "International Land Quality Indexes" University of Minnesota Dept. of Agricultural and Applied Economics Staff Paper P87-10.

Phillips, J. M. 1994. "Farmer Education and Farmer Efficiency: A Meta-Analysis." *Economic Development and Cultural Change*, *43*(1): 149-65.

Political Risk Services. 2005. *International Country Risk Guide*. New York: Political Risk Services.

Pritchett, L. 2001. "Where Has All the Education Gone?" *The World Bank Economic Review*, *15*(3): 367-91.

Rosenzweig, M. R. 1995. "Why Are There Returns to Schooling?" *The American Economic Review*, *85*(2): 153-58.

Schultz, T. P. 1988. "Education Investments and Returns," H. Chenery , and T. N. Srinivasan, *Handbook of Development Economics Vol. 1.* Amsterdam: Elsevier Science, 543-630.

Schultz, T. P. 1999. "Health and Schooling Investments in Africa." *The Journal of Economic Perspectives*, *13*(3): 67-88.

Schultz, T. W. 1975. "The Value of the Ability to Deal with Disequilibria." *Journal of Economic Literature*, *13*(3): 827-46.

Schultz, T. W. 1981. *Investing in People: The Economics of Population Quality*. Berkeley: University of California Press.

Ulubaşoğlu, M. A., and B. A. Cardak. 2007. "International Comparisons of Rural-Urban Educational Attainment: Data and Determinants." *European Economic Review*, *51*(7): 1828-57.

UNDP. 2009. *Human Development Report 2009: Overcoming Barriers: Human Mobility and Development*. New York: United Nations Development Programme.

UNESCO. 2011. "Glossary." (May 12, 2011), *available at* http://www.uis.unesco.org/glossary/

Vollrath, D. 2007. "Land Distribution and International Agricultural Productivity." *American Journal of Agricultural Economics*, *89*(1): 202-16.

Weir, S., and J. Knight. 2004. "Externality Effects of Education: Dynamics of the Adoption and Diffusion of an Innovation in Rural Ethiopia." *Economic Development and Cultural Change*, *53*(1): 93-113.

Welch, F. 1970. "Education in Production." *The Journal of Political Economy*, *78*(1): 35-59.

Williams, R., and V. Breneman. 2009. *Global Agricultural Land Precipitation*. USDA - Economic Research Service, Analysis of the monthly climatic data from the Climate Research Unit (CRU) at the University of East Anglia.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: The MIT Press.

Wouterse, F. "Social Services, Human Capital, and Technical Efficiency of Smallholders in Burkina Faso." *IFPRI Discussion Papers 01068*.

Yang, D. T. 1997. "Education and Off-Farm Work." *Economic Development and Cultural Change*, 613-32.

Yang, D. T., and M. Y. An. 2002. "Human Capital, Entrepreneurship, and Farm Household Earnings." *Journal of Development Economics*, *68*(1): 65-88.

Young, D., and H. Deng. 1999. "The Effects of Education in Early-Stage Agriculture: Some Evidence from China." *Applied Economics*, *31*(11): 1315-23.

Appendix A: Data Description and Sources

a) Variables directly included in the regressions

Variable	Description	Source
(log) Net agricultural production per ha	Net Production (calculated as the total value of all agricultural production after deductions for feed and seed) in 1999-2001 international $ divided by total agricultural area in ha	FAOSTAT
(log) Livestock per ha	Own calculation of the number of cow equivalents by applying the weights from Hayami and Ruttan (1985) to the stocks of live animals obtained from FAOSTAT and then dividing the weighted sum by the total agricultural area in ha	FAOSTAT/ Own calculation
(log) Fertilizer per ha	Total fertilizer consumption in tons divided by total agricultural area in ha	FAOSTAT
(log) Tractors per ha	Number of agricultural tractors in use divided by total agricultural area in ha	FAOSTAT
(log) Workers per ha	Total economically active population in agriculture divided by total agricultural area in ha	FAOSTAT (revision 2006)
Area equipped for irrigation (%)	Share of the total agricultural area that is equipped for irrigation	FAOSTAT
Permanent meadows and pastures (%)	Share of the total agricultural area that is used as permanent meadows and pastures	FAOSTAT
Life expectancy at birth	Average life expectancy at birth in years	World Development Indicators (WDI)
Total fertility rate	Average number of births per woman	World Development Indicators (WDI)
Net enrolment ratio (%)	Net enrolment ratio for primary education (both sexes)	UNESCO Institute for Statistics (UIS)
Gross enrolment ratio (%)	Gross enrolment ratio for primary education (both sexes)	UNESCO Institute for Statistics (UIS)
Adult literacy rate (%)	Adult literacy rate (aged 15 and above - both sexes)	United Nations Development Programme (UNDP)
Total years of schooling	Number of years of total schooling achieved by the average person	Barro and Lee (2010)
Years of primary education	Number of years of primary schooling achieved by the average person	Barro and Lee (2010)
Years of secondary education	Number of years of secondary schooling achieved by the average person	Barro and Lee (2010)
Years of tertiary education	Number of years of tertiary schooling achieved by the average person	Barro and Lee (2010)
Road sector energy consumption	Per capita road sector energy consumption in kilotons of oil equivalents	World Development Indicators (WDI)
Road density	Own calculation of the road density defined as the total road length in km divided by the total land area in 100 sq. km	Own calculation
(log) Precipitation (mm)	Annual average precipitation on agricultural land in mm	Williams and Breneman (2009)
ICRG political risk	The ICRG political risk rating includes a total of 12 weighted variables covering political as well as social attributes (e.g. corruption, bureaucratic quality, external and internal conflict etc.)	International Country Risk Guide. PRS Group (2005).
English legal origin	Dummy variable. Code 1 if the country's legal origin is English common law	La Porta et al. (1999)
French legal origin	Dummy variable. Code 1 if the country's legal origin is French commercial code	La Porta et al. (1999)

German legal origin	Dummy variable. Code 1 if the country's legal origin is German commercial code	La Porta et al. (1999)
Scandinavian legal origin	Dummy variable. Code 1 if the country's legal origin is Scandinavian commercial code	La Porta et al. (1999)
Socialist legal origin	Dummy variable. Code 1 if the country's legal origin is Socialist/Communist law	La Porta et al. (1999)
Peterson land quality index	Land quality index for all agricultural land	Peterson (1987)

b) Variables used only for calculation reasons

Variable	**Description**	**Source**
Total agricultural area	Total agricultural area in ha	FAOSTAT
Road length	Total road length in km	Canning (1998)
Total land area	Total land area in sq. km	World Development Indicators (WDI)
GDP per capita, PPP	GDP per capita in PPP (constant 2005 international $)	Penn World Tables (PWT) 7.0
Rural population (%)	Share of the population living in rural areas (% of total population)	World Development Indicators (WDI)

Appendix B: List of Observations dropped due to the two Exclusion Criteria

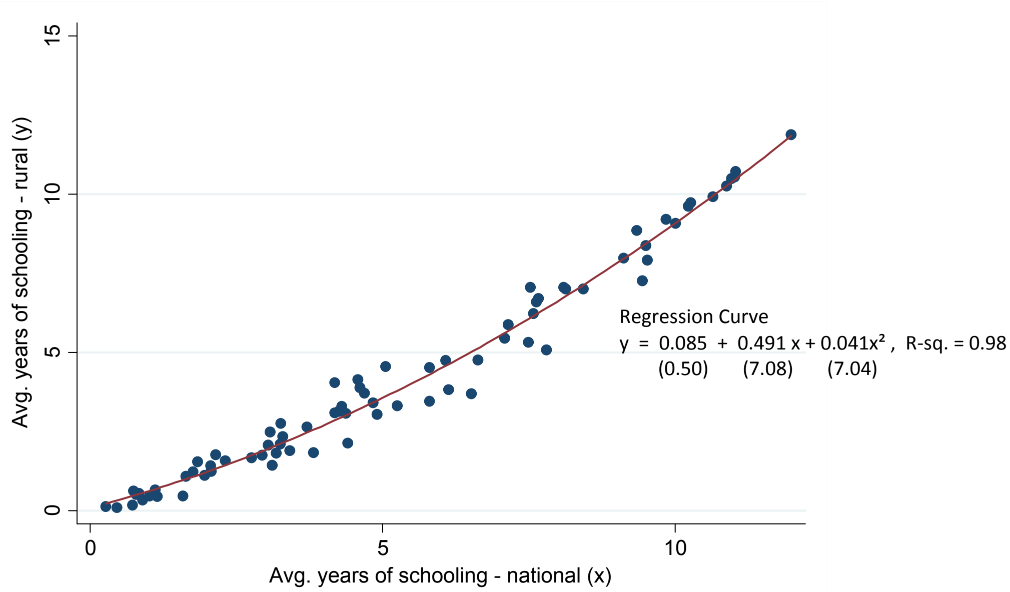
Country	Observations dropped due to the ... natural disaster exclusion rule	armed conflict exclusion rule
Afghanistan	0	12
Albania	1	0
Algeria	0	2
Angola	0	8
Antigua and Barbuda	1	0
Australia	1	0
Bangladesh	1	0
Benin	1	0
Bolivia	1	0
Bosnia and Herzegovina	0	3
Botswana	1	0
Cambodia	1	10
Chad	0	1
Congo	0	1
Djibouti	3	0
Dominican Republic	0	1
El Salvador	1	5
Eritrea	2	2
Fiji	1	0
Gambia	1	0
Ghana	1	0
Guatemala	1	1
Guinea-Bissau	0	11
Guyana	1	0
India	1	0
Iran	1	8
Iraq	0	12
Israel	0	3
Jordan	0	1
Kenya	1	0
Kiribati	1	0
Lao People's Democratic Republic	1	0
Lebanon	0	9
Lesotho	1	0
Liberia	0	1
Libya	0	1
Malawi	1	0
Mauritania	3	0
Mozambique	2	4
Nicaragua	0	6
Niger	1	0
Rwanda	1	0
Sao Tome and Principe	1	0
Senegal	1	0
Sierra Leone	0	1
Somalia	0	3
Sudan	1	0
Swaziland	1	0
Tajikistan	1	2
Uganda	0	6
Vietnam	0	12
Zimbabwe	2	2

Appendix C: List of Countries and Number of Observations

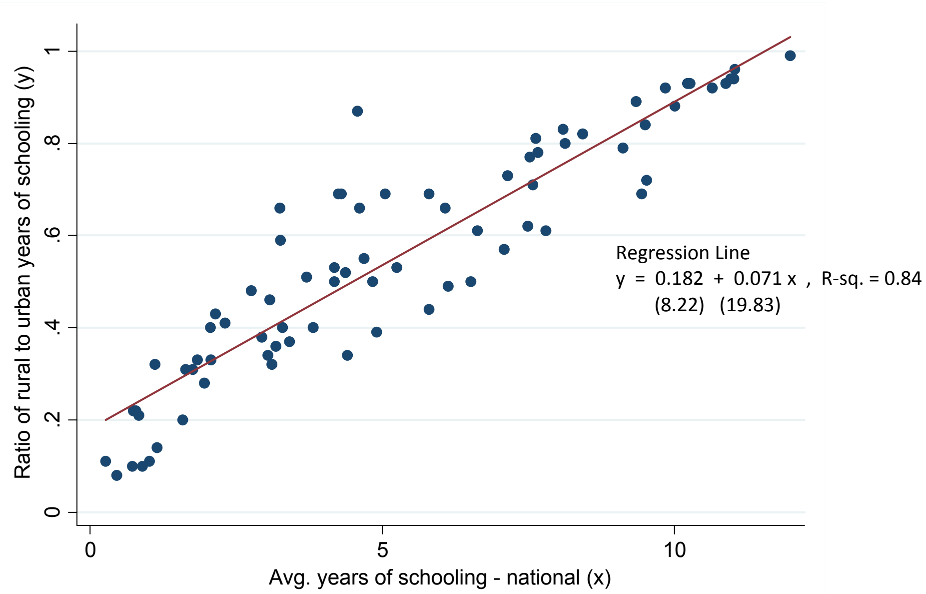
Country	Observations	Country	Observations
Afghanistan	28	Liberia	20
Albania	41	Libya	41
Algeria	40	Lithuania	11
Argentina	42	Malawi	41
Armenia	11	Mali	42
Bangladesh	41	Mauritania	30
Belize	33	Mauritius	24
Benin	38	Mongolia	34
Bolivia	41	Morocco	42
Botswana	41	Mozambique	36
Brazil	42	Myanmar	42
Bulgaria	42	Namibia	5
Burundi	37	Nepal	42
Cambodia	24	Nicaragua	36
Cameroon	42	Niger	38
Central African Republic	23	Pakistan	42
China	42	Panama	42
Colombia	42	Paraguay	42
Congo	31	Peru	42
Costa Rica	42	Philippines	42
Côte d'Ivoire	42	Republic of Moldova	11
Croatia	9	Romania	38
Cuba	42	Russian Federation	11
Cyprus	42	Rwanda	33
Democratic Republic of the Congo	26	Saudi Arabia	42
Dominican Republic	41	Senegal	41
Ecuador	42	Sierra Leone	41
El Salvador	36	South Africa	42
Fiji	41	Sri Lanka	42
Gabon	31	Sudan	41
Gambia	38	Swaziland	41
Ghana	41	Syrian Arab Republic	42
Guatemala	40	Tajikistan	8
Guyana	41	Thailand	42
Haiti	42	Togo	37
Honduras	42	Trinidad and Tobago	42
India	41	Tunisia	24
Indonesia	42	Uganda	35
Iran	33	Ukraine	11
Iraq	32	United Arab Emirates	31
Jamaica	42	United Republic of Tanzania	42
Jordan	41	Uruguay	21
Kazakhstan	11	Venezuela	23
Kenya	41	Vietnam	30
Kyrgyzstan	11	Yemen	37
Laos	37	Zambia	42
Latvia	11	Zimbabwe	36
Lesotho	41		

- **Total of 3,282 observations from 95 countries (basic FE-specification)**

Appendix D: Relationship between the average years of schooling (national) and the average years of schooling (rural) (based on data from Ulubaşoğlu and Cardak (2007))



Appendix E: Relationship between the average years of schooling (national) and the ratio of rural to urban education (based on data from Ulubaşoğlu and Cardak (2007))



Appendix F: Derivation of the formula for s_{rural}

The avg. years of schooling for the whole nation ($s_{national}$) can be written as the weighted average of the years of schooling for urban (s_{urban}) and rural areas (s_{rural}) whereat the weights are the share of the population living in urban and rural surroundings (ω_{urban} and ω_{rural} , respectively).

(3) $$s_{national} = \omega_{urban} s_{urban} + \omega_{rural} s_{rural}$$

In addition, the ratio of rural to urban education (r) is defined as

(4) $\mathrm{r} = \frac{\mathrm{s}_{\mathrm{rural}}}{\mathrm{s}_{\mathrm{urban}}}$ which can be transformed to

(5) $$\mathrm{s}_{\mathrm{urban}} = \frac{\mathrm{s}_{\mathrm{rural}}}{\mathrm{r}}$$

Replacing $\mathrm{s}_{\mathrm{urban}}$ in equation (3) with (5) gives

(6) $$s_{national} = \omega_{urban} \frac{s_{rural}}{r} + \omega_{rural} s_{rural}$$

This expression can be solved for $\mathrm{s}_{\mathrm{rural}}$ yielding

(7) $s_{rural} = \frac{s_{national}}{\frac{\omega_{urban}}{r} + \omega_{rural}}$ q.e.d.

Appendix G: Results of the main regression when using the estimated average years of schooling for the rural population instead of the Barro-Lee measures

	Method 1		Method 2	
	(1)	(2)	(3)	(4)
	FE	FGLS	FE	FGLS
(log) Livestock per ha	0.296***	0.289***	0.275***	0.287***
	(5.282)	(13.483)	(4.854)	(13.382)
(log) Fertilizer per ha	0.061***	0.017***	0.058***	0.014***
	(3.226)	(4.098)	(3.015)	(3.629)
(log) Tractors per ha	0.077***	0.063***	0.077***	0.065***
	(2.835)	(8.205)	(2.848)	(8.564)
(log) Workers per ha	0.217***	0.242***	0.255***	0.270***
	(2.802)	(8.141)	(3.325)	(8.632)
Area equipped for irrigation (%)	0.004*	0.005***	0.004*	0.005***
	(1.672)	(4.764)	(1.938)	(5.212)
Permanent meadows and pastures (%)	-0.010***	-0.008***	-0.010***	-0.008***
	(-2.820)	(-7.208)	(-2.945)	(-6.818)
Life Expectancy at birth	0.013***	0.011***	0.012***	0.011***
	(3.377)	(7.239)	(3.008)	(7.138)
Total years of rural schooling	0.062**	0.025***	0.021**	0.013***
	(2.471)	(2.778)	(2.549)	(4.479)
Road sector energy consumption	0.217*	0.265***	0.221*	0.248***
	(1.830)	(3.626)	(1.815)	(3.400)
(log) Precipitation (mm)	0.027	0.021**	0.028	0.021**
	(1.172)	(2.390)	(1.204)	(2.383)
Constant	5.940***	5.625***	6.092***	5.671***
	(15.569)	(37.703)	(16.924)	(42.664)
Observations	1,685	1,685	1,685	1,685
Number of countries	74	74	74	74
Time fixed effects	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes
ε_{it} autocorrelation	none	AR(1)	none	AR(1)
Wooldridge test statistic[a]	18.13		18.03	
Wooldridge test p-value	0.00		0.00	
R^2	0.913	n.a.	0.919	n.a.

Notes: The dependent variable is the logarithm of the net agricultural production per ha (in intl. $). Robust t-statistics are given in parentheses.
Single asterisk (*) denotes significance at 10%, double asterisk (**) denotes significance at 5%, and triple asterisk (***) denotes significance at the 1% level.
[a] Wooldridge test statistic is distributed as F(1,73) in columns (1) and (3).